tering can be easily taken into account by adding additional complex contributions $V'_{g\parallel}(z)$ and V'_g to (2) and (8) (Yoshioka, 1957; Nagano, 1990) in two- and three-dimensional Bloch-wave theories, respectively.

References

BETHE, H. (1928). *Ann. Phys. (Leipzig)*, **87**, 55–129.
COWLEY, J. M. & MOODIE, A. F. (1957). *Acta Cryst.* **10**, 609–619.
DARWIN, C. G. (1914). *Philos. Mag.* **27**, 675–690.
DOYLE, P. A. & TURNER, P. S. (1967). *Acta Cryst.* A**24**, 390–397.
FIELDS, P. M. & COWLEY, J. M. (1978). *Acta Cryst.* A**34**, 103–112.
FUJIMOTO, F. (1959). *J. Phys. Soc. Jpn*, **14**, 1558–1568.
ICHIMIYA, A. (1983). *Jpn. J. Appl. Phys.* **22**, 176–180.
ISHIZUKA, K. & UYEDA, N. (1977). *Acta Cryst.* A**33**, 740–749.
KIM, H. S. & SHEININ, S. S. (1982). *Phys. Status Solidi B*, **109**, 807–816.
MAGNUS, S. (1954). *Commun. Pure Appl. Math.* **7**, 649–673.
MAKSYM, P. A. (1985). *Surf. Sci.* **149**, 157–174.
MAKSYM, P. A. & BEEBY, J. L. (1981). *Surf. Sci.* **110**, 423–436.
NAGANO, S. (1990). *Phys. Rev. B*, **42**, 7363–7369.
PENDRY, J. B. (1974). *Low-Energy Electron Diffraction.* London: Academic Press.
PENG, L.-M. & WHELAN, M. J. (1990). *Proc. R. Soc. London. Ser. A*, **431**, 111–123.
VAN DYCK, D. (1980). *J. Microsc.* **119**, 141–152.
WANG, Z. L. (1990). *Phys. Rev. B*, **41**, 12818–12837.
WATANABE, K., KIKUCHI, Y., HIRATSUKA, K. & YAMAGUCHI, H. (1990). *Acta Cryst.* A**46**, 94–98.
YOSHIOKA, H. (1957). *J. Phys. Soc. Jpn*, **12**, 618–628.
ZHAO, T. C., POON, H. C. & TONG, S. Y. (1988). *Phys. Rev. B*, **38**, 1172–1182.

Acta Cryst. (1994). A**50**, 741–748

On the Statistical Analysis of Orientation Data

BY NIELS C. KRIEGER LASSEN AND DORTE JUUL JENSEN

Materials Department, Risø National Laboratory, DK-4000 Roskilde, Denmark

AND KNUT CONRADSEN

The Institute of Mathematical Statistics and Operations Research, The Technical University of Denmark, DK-2800 Lyngby, Denmark

(*Received* 18 *February* 1994; *accepted* 22 *April* 1994)

Abstract

The statistical analysis of data in the form of orientations is a relatively new discipline and results from the literature on this subject are not yet widely known outside the statistics community. This paper provides an introduction to and the key references for statistical methods for analysing orientation data. More specifically, the problem of estimating an unknown orientation is considered and results on the precision of such an estimated orientation are described. The calculation of average orientation and dispersion parameters for a sample of orientations is also considered. Finally, procedures for generating and testing for random orientations are described. The methodology is illustrated with crystal orientation data obtained from the analysis of electron back-scattering patterns.

1. Introduction

Orientation data arise naturally in many scientific areas, notably in the earth sciences, astronomy and biology. Within the field of materials science, for example, the development of techniques for measuring local lattice orientations in polycrystalline materials has opened up the way for wholly new types of investigations. These techniques, especially the electron back-scattering pattern (EBSP) technique in scanning electron microscopy and the Kikuchi diffraction technique in transmission electron microscopy, have recently been a major subject at several conferences and workshops (Bunge, 1993, 1994). These modern techniques for measuring local lattice orientations are convenient and rapid in use and the EBSP technique has recently even been fully automated (Wright & Adams, 1992), thus allowing large numbers of orientation data to be collected. Statistical methods for analysing such data, however, have only very rarely been applied. As a result of this, for example, little is known about the precision by which crystal orientations can be determined. Another example of the application of orientation statistics is when the average and dispersion of a sample of orientations is to be determined. For example, it is demonstrated in the final section of this paper that using the arithmetic mean of Euler angles as a measure of an average orientation cannot generally be recommended.

The aim of the present paper is to introduce and present statistical methods for analysing orientation data in a reasonably accessible form that should allow scientists without a statistical background to apply the results. The computations involved may seem complicated for those unfamiliar with matrix calculations and multivariate statistics but it should be noted that these calculations are easily (and rapidly) performed using standard mathematical library routines. It must also be emphasized that many extensions, generalizations and details have of necessity been omitted and the reader is encouraged to study the original publications.

The following sections deal with the problem of estimating an unknown orientation and describing its precision. They also describe how a sample of orientations may be summarized or modelled by a probability distribution. On the basis of this distribution, parameters describing the average and concentration of orientations are introduced. Finally, it is shown how one may test for and conveniently generate random orientations. The techniques are illustrated with orientation data obtained from the analysis of electron back-scattering patterns but are of course generally applicable.

2. Background and basic principles

In classical statistics, the observed data are in the form of counts, real numbers, unrestricted vectors or matrices. The topic of orientation statistics is, on the other hand, concerned with observations that are rotations of space (usually two- or three-dimensional) and the data are therefore restricted matrices as described below. Orientation statistics are closely related to the better established field of directional statistics, which is concerned with directions in space. A recent overview of the statistics of directions is given by Jupp & Mardia (1989) and the books by Mardia (1972), Watson (1983) and Fisher, Lewis & Embleton (1987) are classical references. The following description is limited to directions and orientations in three-dimensional space since this covers most situations of practical importance.

Initially, orientations and their basic properties must be defined in mathematical terms. The orientation of an object in three-space can generally be described by a rigid configuration of three distinguishable directions (see Downs, 1972). These directions are conveniently represented by three 3×1 column vectors $\mathbf{x}_i$, $i = 1, 2, 3$, whose elements are the coordinates of the ith direction measured in some fixed standard (rectangular and right-handed) Cartesian coordinate system, the reference system. Furthermore, it is convenient to let the vectors be of unit length, $\mathbf{x}_i^T\mathbf{x}_i = 1$ and aligned so that $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ form the axes of a standard Cartesian coordinate system $(O,\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)$. If $\mathbf{X}$ is the 3×3 matrix whose columns are $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$, the orientation of the object is rigidly described by $\mathbf{X}$. It is evident that $\mathbf{X}^T\mathbf{X} = \mathbf{X}\mathbf{X}^T = \mathbf{I}$, where $\mathbf{I}$ is the 3×3 identity matrix, and that $\det(\mathbf{X}) = +1$, since $\mathbf{X}$ represents the rotation between two right-handed Cartesian coordinate systems. The collection of 3×3 matrices $\mathbf{X}$ satisfying $\mathbf{X}^T\mathbf{X} = \mathbf{I}$ and $\det(\mathbf{X}) = +1$ is a representation of the group of all proper rotations, the so-called *special orthogonal group* SO(3).

Two fundamental distributions for rotations and directions are briefly introduced in the following. In some respects they can both be regarded as the 'normal distributions' for data consisting of rotations and directions, respectively.

As mentioned above, directions in three-space are conveniently described by unit vectors $\mathbf{x}$ satisfying $\mathbf{x}^T\mathbf{x} = 1$. Thus, the sample space of directional data is the unit sphere S^2 in three-space. The most widely used distribution employed for modelling spherical or directional data is the *Fisher* (or Fisher–von Mises) distribution (Fisher, 1953) with probability density

$$f(\mathbf{x};\mathbf{m},k) = (k/4\pi \sinh k)\exp(k\mathbf{m}^T\mathbf{x}) \qquad (1)$$

with respect to the uniform distribution on S^2, where the mean direction $\mathbf{m}$ is the unit three-vector maximizing (1) and $k \geq 0$ is a concentration parameter. The distribution is rotationally symmetric around $\mathbf{m}$ and becomes successively more concentrated as k approaches infinity. For $k = 0$, $\mathbf{x}$ is uniformly distributed over the unit sphere S^2. Given a sample of directions $\mathbf{x}_i$, the maximum-likelihood estimate $\hat{\mathbf{m}}$ of the mean direction is found by maximizing $\sum\mathbf{m}^T\mathbf{x}_i = \sum\cos\theta_i$, where θ_i is the angle between $\mathbf{m}$ and $\mathbf{x}_i$. For the maximum-likelihood estimate of k, see, for example, Watson (1983).

The Fisher distribution (1) for directional or spherical data can be generalized to describe a distribution for orientations $\mathbf{X}$ on SO(3) (or more generally on Stiefel manifolds) as proposed by Downs (1972). This distribution is now known as the *matrix Fisher* distribution (Khatri & Mardia, 1977) and it has probability density

$$\mathbf{F}(\mathbf{X};\mathbf{M},\mathbf{K}) = \mathbf{a}(\mathbf{K})\exp[\mathrm{tr}(\mathbf{K}\mathbf{M}\mathbf{X}^T)] \qquad (2)$$

with respect to the uniform distribution on SO(3). $\mathbf{F} = \mathbf{K}\mathbf{M}$ is a 3×3 parameter matrix that describes both the mean orientation $\mathbf{M}\in$ SO(3) and the concentration $\mathbf{K}$ of the distribution, where $\mathbf{K}$ is a 3×3 symmetric matrix. $a(\mathbf{K})$ is a normalizing constant that depends only upon the eigenvalues of $\mathbf{K}$. That the matrix Fisher distribution for orientations is a generalization of the Fisher distribution for directional data may be clarified by the following observations: The exponent in (1), $\exp(k\mathbf{m}^T\mathbf{x})$, may be written $\exp[\mathbf{m}^T(k\mathbf{I})\mathbf{x}]$ and a generalization would

then be $\exp[\mathrm{tr}(\mathbf{M}^T\mathbf{K}\mathbf{X})] = \exp[\mathrm{tr}(\mathbf{K}\mathbf{M}\mathbf{X}^T)]$, the exponent of (2). Note that, if $\mathbf{K} = K\mathbf{I}$, $\exp[\mathrm{tr}(\mathbf{K}\mathbf{M}\mathbf{X}^T)] = \exp[K\mathrm{tr}(\mathbf{M}\mathbf{X}^T)]$ and hence the matrix Fisher distribution becomes 'isotropic', *i.e.* depends solely on the orientation 'distance' $\mathrm{tr}(\mathbf{M}\mathbf{X}^T)$. For a sample of orientations $\mathbf{X}_i$, the maximum-likelihood estimate $\hat{\mathbf{M}}$ of the mean orientation is found by maximizing $\sum\mathrm{tr}(\mathbf{M}\mathbf{X}_i^T)$. The form of $a(\mathbf{K})$ is unfortunately extremely complicated and this makes maximum-likelihood estimation of the parameter matrix $\mathbf{K}$ very difficult. However, some simplifications and approximations have been described recently by Bingham, Chang & Richards (1992) and Wood (1993).

3. The estimation of an unknown orientation

As described above, the orientation of an object in three-space can be represented by a matrix $\mathbf{X} \in \mathrm{SO}(3)$, which describes the rotation between two standard Cartesian coordinate systems; one fixed to the object $(\mathbf{O},\hat{\mathbf{u}}_1,\hat{\mathbf{u}}_2,\hat{\mathbf{u}}_3)$ and one fixed to the 'world' $(\mathbf{O},\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2,\hat{\mathbf{v}}_3)$, the reference or world system (shortly denoted U and V, respectively). Since only the rotation of the two frames is considered (not the translation), U and V can be thought to have a common origin O. The orientation of the object with respect to the reference system V is then represented by the matrix $\mathbf{X}$, whose columns $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ contain the coordinates of $\hat{\mathbf{u}}_1$, $\hat{\mathbf{u}}_2$ and $\hat{\mathbf{u}}_3$ measured in V. When crystal orientations are determined from electron back-scattering patterns (see, for example, Randle, 1992), the axes of U are chosen parallel to certain standard directions in the crystal lattice and the axes of V are chosen so that $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ are in the plane of the pattern.

To determine the unknown orientation $\mathbf{X}$, at least two directions (unit vectors) must be measured in both U and V. Let $(\mathbf{u}_i,\mathbf{v}_i)$, $i = 1,\ldots,n$, represent n pairs of 3×1 column vectors of unit length measured in U and V, respectively, so that, without errors, we would have $\mathbf{v}_i = \mathbf{X}\mathbf{u}_i$. In most practical situations, the vectors $\mathbf{u}_i$ and/or $\mathbf{v}_i$ are measured with some error and $\mathbf{X}$ must be estimated from the data using an appropriate measure of the total error. This problem is to some extent similar to the problem of fitting a set of data points to a straight-line model; the linear regression problem. The matrix $\hat{\mathbf{X}}$ minimizing the sum of squared errors defined by

$$\mathrm{SSE}(\mathbf{X}) = \sum_i |\mathbf{v}_i - \mathbf{X}\mathbf{u}_i|^2 = 2n - 2\sum_i \mathbf{v}_i^T\mathbf{X}\mathbf{u}_i \qquad (3)$$

is described as the least-squares estimate of $\mathbf{X}$. The problem of calculating $\hat{\mathbf{X}}$ was first considered and solved by MacKenzie (1957) [though his solution may result in $\det(\mathbf{X}) = -1$] and later refined by Stephens (1979). The least-squares estimate of the unknown orientation $\mathbf{X}$ can be found by performing a singular-value decomposition (for a routine see, for example, Press, Flannery, Teukolsky & Vetterling, 1988) of a 3×3 matrix $\mathbf{A}$:

$$\mathbf{A} = (1/n)\sum_i \mathbf{u}_i\mathbf{v}_i^T = \mathbf{O}_1\Lambda\mathbf{O}_2^T \qquad (4)$$

modified so that $\mathbf{O}_1$, $\mathbf{O}_2 \in \mathrm{SO}(3)$ and $\Lambda = \mathrm{diag}(\lambda_1,\lambda_2,\lambda_3)$ (a diagonal 3×3 matrix) with $\lambda_1 \geq \lambda_2 \geq |\lambda_3|$. If not more than one of the singular values λ_i of $\mathbf{A}$ is equal to zero, the least-squares estimate is uniquely given by

$$\hat{\mathbf{X}} = \mathbf{O}_2\mathbf{O}_1^T. \qquad (5)$$

When only two pairs of vectors $(\mathbf{u}_i,\mathbf{v}_i)$, $i = 1,2$, are available ($n = 2$), $\mathbf{A}$ is singular and the singular value λ_3 is equal to zero but the least-squares estimate can still uniquely be found from (5).

In many practical situations, the vectors $\mathbf{u}_i$ are (or can be assumed to be) known without error. This is, for example, true when crystal orientations are determined from electron back-scattering patterns since $\mathbf{u}_i$ here represents the normal to specific (and known from the process of indexing) crystal planes. Furthermore, it is often reasonable to assume that the $\mathbf{v}_i$'s are rotationally symmetrically distributed around $\mathbf{X}\mathbf{u}_i$ and that the underlying distribution is Fisher $f(\mathbf{x};\mathbf{X}\mathbf{u}_i,k)$. Under these assumptions, it can be shown that the least-squares estimate $\hat{\mathbf{X}}$ found from (5) is also a maximum-likelihood estimate of the unknown orientation $\mathbf{X}$ (Moran, 1976).

The problem of estimating an unknown orientation as described above was named *spherical regression* by Chang (1986), who made a comprehensive study of the statistical properties of $\hat{\mathbf{X}}$. Recent surveys of this subject are given by Watson (1989) and Chang (1993).

4. A confidence region for X

In order to make inferences about the precision of an orientation estimated as above, some confidence regions are described. Such confidence regions can only be described analytically when certain assumptions about the data are made. Under different assumptions about the underlying distribution and/or sample size, approximate confidence regions for $\mathbf{X}$ have been described by Chang (1986), Rivest (1989) and Chang (1989) and a very useful overview is given by Chang (1987).

Assume here that $\mathbf{u}_i$ are known without error and that $\mathbf{v}_i$ are independently Fisher distributed $f(\mathbf{v}_i;\mathbf{X}\mathbf{u}_i,k)$ with common concentration parameter k. If k is large (corresponding to relatively small errors), an approximate confidence region for the orientation $\mathbf{X}$ can be determined as described in the following; for confidence regions obtained using other assumptions,

the reader should consult Chang (1987). In the case of crystal-orientation determination from electron back-scattering patterns, both the Fisher model and the large-k assumption can be shown to be suitable. The Fisher model was verified by applying different goodness-of-fit tests [both so-called Q–Q plots and more formal procedures were used; see, for example, Fisher *et al.* (1987), p. 117] to data sets $(\mathbf{u}_i,\mathbf{v}_i)$ from several different patterns. Data from several patterns are needed to verify the Fisher model since typically only 8–12 bands (data points) are observed in a single pattern and 25 is regarded as the minimum number of data points for a reliable test (Fisher *et al.*, 1987, p. 122). All goodness-of-fit tests showed good agreement with the Fisher model unless gross errors resulting from incorrect indexing were present in the data set. With regard to the large-k assumption, a k value of, say, 100 is usually assumed to be large and since typical values for electron back-scattering patterns are of the order of 3000–20 000, this assumption seems appropriate.

In order to describe a confidence region for estimated orientations, a mathematically convenient parameterization of SO(3) must be introduced [for a discussion on this subject see Chang (1993)]. The so-called exponential parameterization turns out to be very useful for describing small rotations such as those in a confidence region. The parameter vector $\mathbf{h}\in R^3$ is mapped by Φ to an orientation $\Phi(\mathbf{h})\in$ SO(3), which represents right-hand-rule rotation of $|\mathbf{h}|$ rad around the axis $\mathbf{h}/|\mathbf{h}|$. If one denotes the rotation angle $\omega=|\mathbf{h}|$ and the rotation axis $\mathbf{t}=[t_1t_2t_3]^T=\mathbf{h}/|\mathbf{h}|$, then the corresponding orientation matrix can be found from

$$\mathbf{\Phi(h)}=\mathbf{I}+\sin\omega\begin{pmatrix}0&-t_3&t_2\\t_3&0&-t_1\\-t_2&t_1&0\end{pmatrix}+(1-\cos\omega)\begin{pmatrix}0&-t_3&t_2\\t_3&0&-t_1\\-t_2&t_1&0\end{pmatrix}^2. \quad (6)$$

Conversely, if an orientation $\mathbf{X}$ is given, ω and $\mathbf{t}$ (and $\mathbf{h}=\omega\mathbf{t}$) can be found using

$$1+2\cos\omega=\mathrm{tr}\,(\mathbf{X})$$

$$\mathbf{X}-\mathbf{X}^T=2\sin\omega\begin{pmatrix}0&-t_3&t_2\\t_3&0&-t_1\\-t_2&t_1&0\end{pmatrix}. \quad (7)$$

It should be emphasized that the exponential parameterization presented above is not only convenient for describing small rotations but is a generally applicable and very useful parameterization.

In order to describe a confidence for $\mathbf{X}$, the following symmetric 3×3 matrix $\mathbf{\Sigma}$ must first be calculated:

$$\mathbf{\Sigma}=(1/n)\sum_i\mathbf{u}_i\mathbf{u}_i^T. \quad (8)$$

This matrix summarizes information about the position of the points corresponding to $\mathbf{u}_i$. Let $\sigma_1\geq\sigma_2\geq\sigma_3\geq0$ be its eigenvalues satisfying $\sigma_1+\sigma_2+\sigma_3=1$ and let $\mathbf{z}_1$, $\mathbf{z}_2$ and $\mathbf{z}_3$ be the corresponding eigenvectors. The eigenvector $\mathbf{z}_1$ will lie in the centre of the points $\mathbf{u}_i$, and the plane defined by $\mathbf{z}_1$ and $\mathbf{z}_2$ will intersect the unit sphere in the great circle that best fits $\mathbf{u}_i$. If $\sigma_i=1/3$, the $\mathbf{u}_i$'s are uniformly distributed on the unit sphere.

If k is known, an approximate $(1-\alpha)$ confidence region for $\mathbf{X}$ contains all orientations $\Phi(\mathbf{h})\hat{\mathbf{X}}$ with $\mathbf{h}$ satisfying

$$\mathbf{h}^T(\mathbf{I}-\hat{\mathbf{X}}\mathbf{\Sigma}\hat{\mathbf{X}}^T)\mathbf{h}<(1/nk)\chi^2_{1-\alpha}(3), \quad (9)$$

where $\chi^2_{1-\alpha}(3)$ is the $(1-\alpha)$ percentage point of the χ^2 distribution with three degrees of freedom. The confidence region is thus made apparent by following the estimated orientation $\hat{\mathbf{X}}$ by a small rotation $\Phi(\mathbf{h})$, where $\mathbf{h}$ must satisfy (9). The region of $\mathbf{h}$ satisfying (9) is an ellipsoid with axes $\hat{\mathbf{X}}\mathbf{z}_1$, $\hat{\mathbf{X}}\mathbf{z}_2$ and $\hat{\mathbf{X}}\mathbf{z}_3$ and corresponding axis lengths ω_i given by

$$\omega_i=[(1/nk)\chi^2_{1-\alpha}(3)/(1-\sigma_i)]^{1/2}. \quad (10)$$

The axis lengths are denoted ω_1, ω_2 and ω_3 to emphasize that they represent angles of rotation. The largest rotation that is allowed within the confidence region for $\mathbf{X}$ is thus ω_1 in the direction of $\hat{\mathbf{X}}\mathbf{z}_1$.

In most practical situations, the concentration parameter of the Fisher distribution k is unknown and must be estimated from the data. For large k, a good approximation to the maximum-likelihood estimate of k is (Watson, 1983)

$$\hat{k}=2n/\sum_i|\mathbf{v}_i-\hat{\mathbf{X}}\mathbf{u}_i|^2. \quad (11)$$

When k has been estimated from (11), the approximate $(1-\alpha)$ confidence region for $\mathbf{X}$ contains all orientations $\Phi(\mathbf{h})\hat{\mathbf{X}}$ with $\mathbf{h}$ satisfying

$$\mathbf{h}^T(\mathbf{I}-\hat{\mathbf{X}}\mathbf{\Sigma}\hat{\mathbf{X}}^T)\mathbf{h}<(3/n\hat{k})F_{1-\alpha}(3,2n-3), \quad (12)$$

where $F_{1-\alpha}(3,2n-3)$ is the $(1-\alpha)$ percentage point of the F distribution with $(3,2n-3)$ degrees of freedom. For the case of unknown k, (10) is modified by replacing $\chi^2_{1-\alpha}(3)$ with $3F_{1-\alpha}(3,2n-3)$.

5. The average and concentration of orientations

Consider now the problem of summarizing a given sample of orientations $\mathbf{X}_i\in$ SO(3), $i=1,\ldots,n$, by fitting it to a model that depends on a few adjustable parameters. The model considered here is the matrix Fisher distribution $a(\mathbf{K})\exp[\mathrm{tr}\,(\mathbf{KMX}^T)]$ so the

sample is summarized by the mean orientation **M** and the concentration matrix **K**.

The 3×3 parameter matrix $\mathbf{F}=\mathbf{KM}$ is said to decompose into its polar form, where **M** and **K** are known as the polar and elliptical components of **F**, respectively. **K** is a 3×3 symmetric matrix assumed to be definite (positive or definite) and **M** belongs to SO(3). The matrix **F** also has a singular-value decomposition defined by

$$\mathbf{F}=\mathbf{\Delta D}_\phi\mathbf{\Gamma} \tag{13}$$

and modified so that $\mathbf{\Delta}, \mathbf{\Gamma} \in \mathrm{SO(3)}$ and $\mathbf{D}_\phi = \mathrm{diag}(\phi_1,\phi_2,\phi_3)$ with $\phi_1 \geq \phi_2 \geq |\phi_3|$. The polar form is related to this singular-value decomposition through

$$\mathbf{M}=\mathbf{\Delta\Gamma}, \quad \mathbf{K}=\mathbf{\Delta D}_\phi\mathbf{\Delta}^T. \tag{14}$$

It can be shown (Downs, 1972; Prentice, 1986) that the matrix Fisher distribution $a(\mathbf{K})\exp[\mathrm{tr}(\mathbf{KMX}^T)]$ has a maximum at $\mathbf{X}=\mathbf{M}$ if $\phi_3>0$ and a minimum at $\mathbf{X}=\mathbf{M}$ if $\phi_3<0$. Furthermore, the clustering of **X** around **M** increases with increasing values of ϕ_i. It is therefore customary to describe **M** as the mean orientation of **X** and **K** (or $\mathbf{D}_\phi$) as the concentration matrix. If $\mathbf{K}=\mathbf{0}$, then **X** is uniformly distributed over SO(3).

For a sample of orientations $\mathbf{X}_i$, the maximum-likelihood estimate $\hat{\mathbf{M}}$ of the mean orientation is found by maximizing $\sum\mathrm{tr}(\mathbf{MX}_i^T)$. Since $\mathrm{tr}(\mathbf{MX}_i^T) = 2\cos\omega_i + 1$, where ω_i is the rotation angle between **M** and $\mathbf{X}_i$, the mean orientation is found by maximizing $\sum\cos\omega_i$. This natural definition of a mean orientation is similar to the definition of a mean direction **m** in the Fisher distribution.

A procedure for calculation of the maximum-likelihood estimate $\hat{\mathbf{M}}$ of the mean orientation **M** is now described. First, the maximum-likelihood estimates of $\mathbf{\Delta}$ and $\mathbf{\Gamma}$ must be found from a singular-value decomposition of the arithmetic mean of the orientations $\mathbf{X}_i$:

$$\overline{\mathbf{X}}=(1/n)\sum_{i=1}^{n}\mathbf{X}_i=\mathbf{\Delta D}_g\mathbf{\Gamma}, \tag{15}$$

modified as in (13) so that $\mathbf{D}_g=\mathrm{diag}(g_1,g_2,g_3)$ with $g_1\geq g_2\geq|g_3|$. The maximum-likelihood estimate of the mean orientation is then $\hat{\mathbf{M}}=\mathbf{\Delta\Gamma}$ (Khatri & Mardia, 1977).

A similar procedure for calculation of the maximum-likelihood estimate of the concentration matrix **K** (or $\mathbf{D}_\phi$) is not easily described, owing to the complicated form of the normalizing constant $a(\mathbf{K})$ (Khatri & Mardia, 1977). However, a procedure that simplifies the calculations of $a(\mathbf{K})$ has been described recently by Wood (1993).

The expectation value of **X** has the polar form $E(\mathbf{X})=\mathbf{HM}$, where the elliptical component **H** is described as the precision matrix (Downs, 1972). The closer **H** is to the identity matrix, the closer the clustering of **X** around **M**. The maximum-likelihood estimate of **H** can be found from the singular-value decomposition (15) as $\hat{\mathbf{H}}=\mathbf{\Delta D}_g\mathbf{\Delta}^T$.

Again, owing to the complicated form of the normalizing constant in the matrix Fisher distribution, tests and confidence regions for the mean orientation **M** and the concentration matrix **K** are not easily described. However, for orientations on SO(3), statistical methods based on moments (without assumptions about an underlying distribution) have been developed by Prentice (1984, 1986, 1989). It is beyond the scope of this paper to give details of these very useful methods and the reader is referred to the original papers.

6. Generating and testing for random orientations

It is well known that orientations in SO(3) can be represented by normalized quaternions (Altmann, 1986). A quaternion Q is an object consisting of a scalar q_0 and a three-dimensional vector $\mathbf{q}=(q_1q_2q_3)^T$: $Q=[q_0;\mathbf{q}]$. If $\mathbf{q}=\mathbf{0}$, Q is a *real* quaternion; if $q_0=0$, Q is a *pure* quaternion; if $q_0=0$ and $|\mathbf{q}|=1$, Q is known as a *unit* quaternion. The norm $|Q|$ of Q is defined by $|Q|^2=q_0^2+|\mathbf{q}|^2=q_0^2+q_1^2+q_2^2+q_3^2$ and a quaternion with unit norm $|Q|=1$ is called *normalized*. In the literature, the normalized quaternion is often confused with the unit quaternion but here we use the notation of Altmann (1986) so that orientations are parameterized by normalized and not unit quaternions. Let $\mathbf{x}=(x_0x_1x_2x_3)^T$ be a four-dimensional vector of unit length representing the normalized quaternion $[x_0;(x_1x_2x_3)^T]$ and let H_4 denote the unit sphere in four-dimensional space with antipodes identified ($+\mathbf{x}$ and $-\mathbf{x}$ are identified). It may then be shown that, if **x** is uniformly distributed on H_4, the corresponding 3×3 orientation matrix $\mathbf{X}(\mathbf{x})\in\mathrm{SO(3)}$ defined by

$$\mathbf{X}(\mathbf{x})=\mathbf{X}(-\mathbf{x})=\begin{pmatrix} x_0^2+x_1^2-x_2^2-x_3^2 & 2(x_1x_2-x_0x_3) & 2(x_1x_3+x_0x_2) \\ 2(x_1x_2+x_0x_3) & x_0^2-x_1^2+x_2^2-x_3^2 & 2(x_2x_3-x_0x_1) \\ 2(x_1x_3-x_0x_2) & 2(x_2x_3+x_0x_1) & x_0^2-x_1^2-x_2^2+x_3^2 \end{pmatrix} \tag{16}$$

is uniformly distributed on SO(3) (Moran, 1976). This elegant correspondence between unsigned directions (axes) in four-space and orientations in three-space provides an easy way to construct random orientations:

(1) Let n_i, $i=0,\ldots,3$, be independent simulations from the standard normal $N(0,1)$ distribution; $n_i\in[-\infty;\infty]$.

(2) Let

$$\mathbf{x}_i=n_i/(n_0^2+n_1^2+n_2^2+n_3^2)^{1/2}. \tag{17}$$

(3) Use the vector $\mathbf{x} = (x_0 x_1 x_2 x_3)^T$ in (16) to obtain a random orientation $\mathbf{X}$.

The first to exploit normalized quaternions in orientation statistics was Mɔran (1976). Later, Prentice (1978, 1986) used the quaternion parameterization of SO(3) to establish a one-to-one relationship between the matrix Fisher distribution on SO(3) and the so-called Bingham distribution on H_4 (Bingham, 1974). It may be shown (Prentice, 1986) that the average orientation matrix $\mathbf{M}$ introduced above in connection with the matrix Fisher distribution can be determined in a simple way using normalized quaternions. If a sample of orientations $\mathbf{X}_i$ is represented by their corresponding quaternions $\pm\mathbf{x}_i$ (with the sign chosen so that $x_0 \geq 0$), the quaternion $\mathbf{m}$ representing $\mathbf{M}$ can be found from the arithmetic mean $\bar{\mathbf{x}} = n^{-1}\sum\mathbf{x}_i$ as $\mathbf{m} = \bar{\mathbf{x}}/|\bar{\mathbf{x}}|$. The authors are grateful to a referee for reminding us of this simple approach to orientation averaging.

Finally, consider the problem of testing samples of orientations $\mathbf{X}_i$ on SO(3) for uniformity. For a detailed discussion on the subject, the reader is referred to Prentice (1978). A likelihood ratio test for uniformity on SO(3) against the alternative of a matrix Fisher distribution (with $\mathbf{K} \neq \mathbf{0}$) uses the generalized Rayleigh statistic

$$\mathbf{R} = 3n[\mathrm{tr}\,(\bar{\mathbf{X}}^T\bar{\mathbf{X}})], \tag{18}$$

where $\bar{\mathbf{X}}$ is the arithmetic mean of the n $\mathbf{X}$'s and R is asymptotically distributed as χ^2_9 (the χ^2 distribution with nine degrees of freedom) under the null hypothesis. The hypothesis is rejected when R is large compared with the percentile values for the χ^2_9 distribution.

7. Applications

For the purpose of illustration, some of the techniques described above are applied to crystal-orientation data determined through the analysis of electron back-scattering patterns obtained in a scanning electron microscope (see, for example, Randle, 1992). An electron back-scattering pattern contains the traces of certain crystallographic planes in the crystal lattice from where the pattern originates. The directions of the normals to these planes are measured in two coordinate systems V and U, where V is fixed to the pattern and serves as the reference frame and U is fixed to certain standard directions in the crystal. The outcome of the analysis of such a pattern is n pairs of unit vectors $(\mathbf{u}_i,\mathbf{v}_i)$, $i = 1,\ldots,n$, which describe n different directions measured in U and V. As an example, Table 1 lists the coordinates of eight pairs of vectors.

The object is to estimate the orientation of the crystal $\mathbf{X}$, measured with respect to the reference frame V [normally the orientation of a crystal is measured with respect to a coordinate system W fixed to the sample but since the rotation between V and W is constant (it does not change from crystal to crystal) this extra change of coordinates is of no interest here]. From (4), the matrix $\mathbf{A}$ and its singular-value decomposition is found to be

Table 1. *The directions of the normals to eight crystal planes measured in a coordinate system fixed to the crystal* $(\mathbf{u}_i)$ *and a coordinate system fixed to the recorded diffraction pattern* $(\mathbf{v}_i)$

$\mathbf{u}_i^T$	$\mathbf{v}_i^T$
$(0,-1,0)/1^{1/2}$	(−0.941737, −0.005831, 0.336299)
$(-1,1,1)/3^{1/2}$	(0.610173, 0.788681, 0.075313)
$(-1,-1,1)/3^{1/2}$	(−0.445052, 0.772446, 0.453051)
$(-1,0,1)/2^{1/2}$	(0.105818, 0.944008, 0.312491)
$(1,-3,-1)/11^{1/2}$	(−0.891762, −0.421736, 0.164007)
$(1,1,0)/2^{1/2}$	(0.750393, −0.659185, 0.048848)
$(0,1,10/2^{1/2}$	(0.874536, 0.308446, 0.374228)
$(1,3,1)/11^{1/2}$	(0.986600, −0.138188, 0.086739)

$$\mathbf{A} = \begin{pmatrix} 0.048631 & -0.275471 & -0.051984 \\ 0.549877 & 0.002959 & -0.040640 \\ 0.169362 & 0.234054 & 0.095917 \end{pmatrix}$$

$$\mathbf{O}_1 = \begin{pmatrix} -0.021560 & 0.760173 & 0.649363 \\ -0.940079 & 0.205650 & -0.271954 \\ -0.340274 & -0.616316 & 0.710189 \end{pmatrix}$$

$$\Lambda = \mathrm{diag}\,(0.580703, 0.371720, 0.047425)$$

$$\mathbf{O}_2 = \begin{pmatrix} -0.991221 & 0.122859 & 0.048847 \\ -0.131711 & -0.949770 & -0.283882 \\ 0.011516 & -0.287823 & 0.957614 \end{pmatrix}.$$

The least-squares estimate of the crystal orientation is then found from (5):

$$\hat{\mathbf{X}} = \begin{pmatrix} 0.146484 & 0.943808 & 0.296257 \\ -0.903492 & 0.005702 & 0.428566 \\ 0.402795 & -0.330444 & 0.853559 \end{pmatrix}.$$

The most popular and widespread parameterization of SO(3) is by the so-called Euler angles $(\varphi_1,\varphi,\varphi_2)$ (see, for example, Randle, 1992). In Euler angles, $\hat{\mathbf{X}}$ corresponds to $(\varphi_1,\varphi,\varphi_2) = (50.635°, 31.399°, 34.655°)$.

In order to describe a confidence region for $\mathbf{X}$, k in the Fisher distribution (1) must first be estimated. Using (11), $\hat{k} = 6586.38$ is obtained. The next step is to calculate $\boldsymbol{\Sigma}$ from (8):

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0.231061 & 0.062500 & -0.145833 \\ 0.062500 & 0.537879 & 0.130682 \\ -0.145833 & 0.130682 & 0.231061 \end{pmatrix}$$

and then calculate its eigenvalues $(\sigma_1,\sigma_2,\sigma_3) = (0.586227,0.367086,0.046687)$ and the corresponding eigenvectors $\mathbf{z}_1 = (0.028199,0.941812, 0.334956)^T$, $\mathbf{z}_2 = (-0.757701,0.198423,0.621705)^T$ and $\mathbf{z}_3 = (0.651992,-0.271328,0.708016)^T$. By conversion of points $\mathbf{h}$ on the ellipsoid defined by (12) to Euler angles, a plot of the 95% confidence region for $\mathbf{X}$ in Euler space can be obtained. A scatter plot of points (represented by small spheres) on the border of this region in Euler space is given in Fig. 1. The largest rotation away from $\hat{\mathbf{X}}$ contained in the 95% confidence region (12) is

$$\omega_1 = \{[3/8(6586.38)]F_{0.95}(3,13)(1-0.586227)\}^{1/2} = 1.24°$$

around the axis $\hat{\mathbf{X}}\mathbf{z}_1 = (0.992253,0.123443, -0.013952)^T$. Equation (12) can be used for testing whether, for example, $\mathbf{A}$ defined as

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

belongs to the 95% confidence region. From $\Phi(\mathbf{h})\hat{\mathbf{X}} = \mathbf{A}$, we get $\Phi(\mathbf{h}) = \mathbf{A}\hat{\mathbf{X}}^T$ and, from (7), $\mathbf{h} = (-0.437730,0.329972,0.080129)^T$. The left-hand side of (12) then gives 0.1606 whereas the right-hand side gives $(3\times 3.41)/(8\times 6586.38) = 0.00019$, so $\mathbf{A}$ is clearly not within the confidence region.

The orientation of another crystal was then calculated ten times as described above, using different sets of crystal planes and therefore different sets of vector pairs $(\mathbf{u}_i,\mathbf{v}_i)$. The results are given in Table 2 using Euler angles. Note that the scattering in the angles φ_1 and φ_2 is quite large because the orientations are close to the singularity in Euler space at $\phi = 0°$.

Table 2. *The orientation of a crystal measured ten times using different sets of crystal planes*

The orientations are represented by Euler angles (°).

φ_1	ϕ	φ_2
168.983	2.383	104.663
187.192	2.067	87.158
172.207	2.350	100.783
181.485	2.152	92.230
167.942	1.920	106.925
175.477	2.827	97.737
152.158	1.768	120.879
167.776	1.991	104.736
164.327	2.232	107.874
173.900	2.241	99.795

The aim is now to determine estimates of the average orientation $\mathbf{M}$ and the parameters $\mathbf{D}_\phi$ describing the scattering of this sample of orientations (assuming the matrix Fisher distribution is a suitable model for the data). From a singular-value decomposition of the arithmetic mean as described in (15),

$$\overline{\mathbf{X}} = \begin{pmatrix} 0.059793 & -0.997418 & 0.037063 \\ 0.998094 & 0.059556 & -0.007753 \\ 0.005534 & 0.037428 & 0.999256 \end{pmatrix}$$

$$\mathbf{\Delta} = \begin{pmatrix} -0.210822 & -0.474136 & 0.854838 \\ -0.019197 & 0.876333 & 0.481324 \\ -0.977336 & 0.085063 & -0.193853 \end{pmatrix}$$

$$\mathbf{D}_g = \mathrm{diag}(1-2.444720\times 10^{-5}, 1-9.823480\times 10^{-5}, 1-1.108722\times 10^{-4})$$

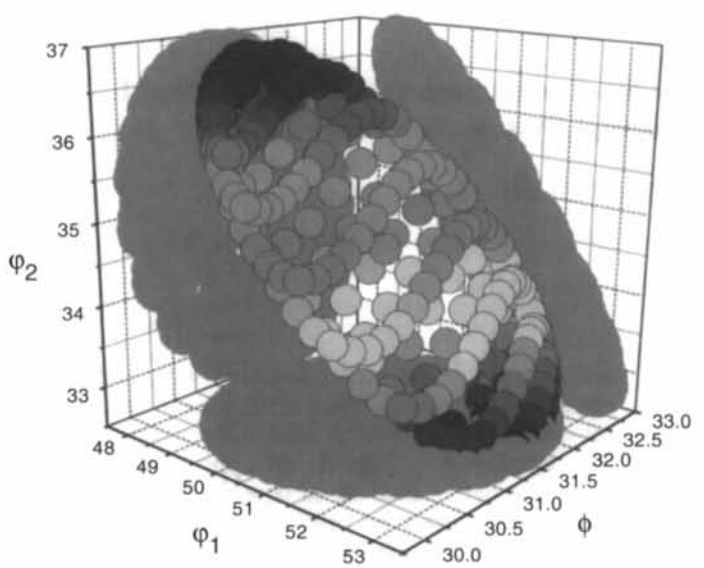

Fig. 1. The 95% confidence region for the crystal orientation determined from the data in Table 1. The small spheres represents points on the boarder of this region in Euler space.

$$\mathbf{\Gamma} = \begin{pmatrix} -0.037176 & 0.172559 & -0.984297 \\ 0.846866 & 0.528338 & 0.060639 \\ 0.530506 & -0.831313 & -0.165776 \end{pmatrix}.$$

A maximum-likelihood estimate of the mean orientation $\mathbf{M}$ in the matrix Fisher distribution is then $\hat{\mathbf{M}} = \mathbf{\Delta\Gamma}$, which expressed in Euler angles is $(\varphi_1,\phi,\varphi_2) = (171.598°,2.169°,101.824°)$. This mean orientation is of course different from the result that would have been obtained by simply taking the arithmetic mean of the Euler angles. Such a simple approach to orientation averaging is not recommended. By the approach of Wood (1993) and with an approximation described by Mardia & Zemroch (1977), the concentration parameter is found as $\mathbf{D}_\phi \simeq \text{diag}(5.311\times10^4, 3.156\times10^4, 2.615\times10^4)$. The large values for ϕ_i indicate, as expected, a large clustering of the $\mathbf{X}$'s.

Consider the problem of testing whether the $\mathbf{X}$'s can be assumed to be uniformly distributed on SO(3). In this case, the answer is clearly no but a formal test for the hypothesis of uniformity can be obtained using the statistic R defined by (18). For the data of Table 2, $R = 3\times10\times2.99953 = 89.986$, which should be compared to χ^2_9. The hypothesis is clearly rejected.

The authors express their appreciation to the referees for their careful reading of the manuscript and helpful suggestions.

References

ALTMANN, S. L. (1986). *Rotations, Quaternions and Double Groups*. Oxford: Clarendon Press.

BINGHAM, C. (1974). *Ann. Stat.* **2**, 1201–1225.

BINGHAM, C., CHANG, T. & RICHARDS, D. (1992). *J. Multivar. Anal.* **41**, 314–337.

BUNGE, H. J. (1993). Editor. *Proc. Microscale Textures of Materials. Textures Microstruct.* **20**, 1–242.

BUNGE, H. J. (1994). Editor. *Proc. ICOTOM*10. In the press.

CHANG, T. (1986). *Ann. Stat.* **14**, 907–924.

CHANG, T. (1987). *J. Geophys. Res.* **B92**, 6319–6329.

CHANG, T. (1989). *Ann. Stat.* **17**, 293–306.

CHANG, T. (1993). *Int. Stat. Rev.* **61**, 299–316.

DOWNS, T. D. (1972). *Biometrika*, **59**, 665–676.

FISHER, N. I., LEWIS, T. & EMBLETON, B. J. J. (1987). *Statistical Analysis of Spherical Data*. Cambridge Univ. Press.

FISHER, R. A. (1953). *Proc. R. Soc. London Ser. A*, **217**, 295–305.

JUPP, P. E. & MARDIA, K. V. (1989). *Int. Stat. Rev.* **57**, 261–294.

KHATRI, C. G. & MARDIA, K. V. (1977). *J. R. Stat. Soc. B*, **39**, 95–106.

MACKENZIE, J. K. (1957). *Acta Cryst.* **10**, 61–62.

MARDIA, K. V. (1972). *Statistics for Directional Data*. London: Academic Press.

MARDIA, K. V. & ZEMROCH, P. J. (1977). *J. Stat. Comput. Simul.* **6**, 29–34.

MORAN, P. A. P. (1976). *Perspectives in Probability and Statistics*, edited by J. GANI, pp. 295–301. Orlando: Academic Press.

PRENTICE, M. J. (1978). *Ann. Stat.* **6**, 169–176.

PRENTICE, M. J. (1984). *Biometrika*, **71**, 147–154.

PRENTICE, M. J. (1986). *J. R. Stat. Soc. B*, **48**, 214–222.

PRENTICE, M. J. (1989). *J. R. Stat. Soc. B*, **51**, 241–248.

PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. & VETTERLING, W. T. (1988). *Numerical Recipes: the Art of Scientific Computing (C version)*. Cambridge Univ. Press.

RANDLE, V. (1992). *Microtexture Determination and its Applications*. London: The Institute of Materials.

RIVEST, L. (1989). *Ann. Stat.* **17**, 307–317.

STEPHENS, M. A. (1979). *Biometrika*, **66**, 41–48.

WATSON, G. S. (1983). *Statistics on Spheres*. New York: Wiley.

WATSON, G. S. (1989). *Probability Measures on Groups* IX, *Lecture Notes in Mathematics*, edited by H. HEYER, pp. 398–413. Berlin: Springer-Verlag.

WOOD, A. T. A. (1993). *Aust. J. Stat.* **35**, 69–79.

WRIGHT, S. I. & ADAMS, R. L. (1992). *Metall. Trans.* **A23**, 759–767.

Acta Cryst. (1994). A**50**, 748–753

STEP – a Trial-and-Error Procedure for Crystal Structure Determination. I. A Description of the Program System *SYSTEM*90

BY HOU YONGGENG, GAO MIN, LI LIPU AND HOU PEIXIN

Institute of Chemistry, Academia Sinica, Beijing 10080, *People's Republic of China*

(*Received* 19 *January* 1994; *accepted* 25 *April* 1994)

Abstract

This paper describes a new trial-and-error direct-methods procedure called *STEP*. A set of strong reflections, sufficient to solve the structure, is divided into a hierarchy of smaller soluble and connected subsystems. Within each subsystem, the reflections are required to be well connected with each other, given that the phases of all reflections in the previous subsystems are known. A trial-and-error procedure is then employed to provide an approximate solution to an overdetermined set of equations. Subsequently, phases are refined by one of two available tangent formulae and then assessed for plausibility by figures